DOCUMENT RESUME

ED 347 254 UD 028 757

TITLE	Schoolwide Project Evaluations: Workshop Guide.
INSTITUTION	Chapter 1 Technical Assistance Center, Denver, CO. Region E.; RMC Research Corp., Denver, CO.
SPONS AGENCY	Office of Elementary and Secondary Education (ED), Washington, DC. Compensatory Education Programs.
REPORT NO	TAC-B-230
PUB DATE	91
NOTE	67p.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Accountability; Comparative Analysis; *Compensatory Education; Data Collection; Elementary Secondary Education; Evaluation Criteria; *Federal Programs; Federal Regulation; Leaders Guides; Middle Schools; *Program Evaluation; Program Implementation; Self Evaluation (Groups); Transparencies; *Workshops
IDENTIFIERS	Hawkins Stafford Act 1988

ABSTRACT

This publication is a guide with the materials necessary for leading a workshop session on Chapter 1 schoolwide project evaluations aimed at meeting federal accountability requirements. As the packet points out, elementary school, middle school, and secondary school projects differ from the traditional Chapter 1 delivery models and as a consequence are required to demonstrate, through these evaluations, that the schoolwide approach is more effective. This workshop teaches participants about the special evaluation requirements for Chapter 1 schoolwide projects. A list of the transparencies and handouts used in he workshop as well as an outline of the 60-75 minute session are included. The actual guide to the workshop covers the following areas: (1) workshop goals; (2) an overview; (3) options for demonstrating effectiveness (same-school comparison and other-school comparison); (4) meeting accountability requirements (subject areas, achievement measures, comparisons, and grade levels); (5) other school comparisons (description and pros and cons); (6) same school comparison (description and baseline score); (7) choosing an approach; and (8) session evaluation. Throughout the guide, graphics indicate when to use the 14 specially designed transparencies. A handout consists of a 16-page paper, "Meeting the Accountability Requirement for Schoolwide Projects" (Alan Davis and Mary R. Quilling). (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

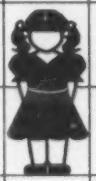
TAC B-230

UD

ED347254

WORKSHOP GUIDE

SCHOOLWIDE PROJECT EVALUATIONS



Includes:

1. PRESENTER'S GUIDE
2. TRANSPARENCIES
3. HANDOUTS
4. RESOURCE MATERIALS

Region E Region 5

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

prepared by:

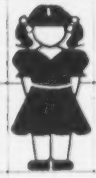
RMC RESEARCH CORPORATION
1512 Larimer Street, Suite 540
Denver, Colorado
80202

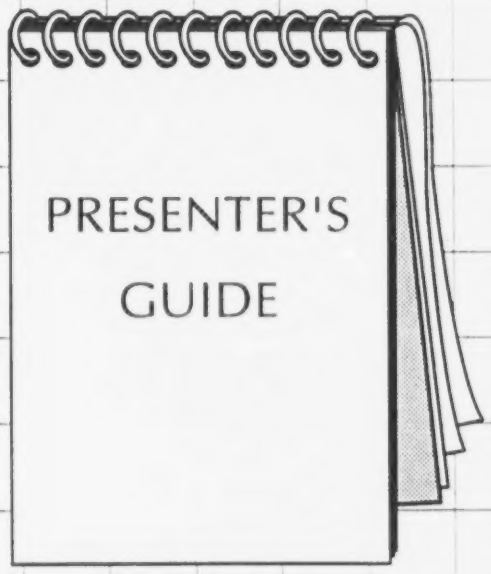
UD 028 75-7

2 BEST COPY AVAILABLE

WORKSHOP GUIDE

SCHOOLWIDE PROJECT EVALUATIONS





prepared by:

RMC RESEARCH CORPORATION
1512 Larimer Street, Suite 540
Denver, Colorado
80202

CHAPTER 1 SCHOOLWIDE PROJECT EVALUATIONS: MEETING THE ACCOUNTABILITY REQUIREMENTS

Transparencies used in this workshop:

T - 1	Schoolwide Projects Must Be More Effective
T - 2	Options for Demonstrating Effectiveness
T - 3	Key Questions
T - 4	Subject Areas
T - 5	Achievement Measures
T - 6	Comparisons
T - 7	Grade Levels
T - 8	Other-School Comparison: Description
T - 9	Other-School Comparison: Pros and Cons
T - 10	Same-School Comparison: Description
T - 11	Same-School Comparison: Compute Baseline Score/Average Gains
T - 12	Same-School Comparison: Compare Baseline Score/Choices
T - 13	Which Approach Works Best?
T - 14	When Neither Approach is Feasible

Handouts used in this workshop:

H - 1	*(optional)* Alan Davis and Mary Quilling, "Meeting the Accountability Requirements for Schoolwide Projects"

Outline

Workshop Goal

Overview T-1

Options for Demonstrating Effectiveness T-2

Meeting Accountability Requirements T-3
- Subject Areas T-4
- Achievement Measures T-5
- Comparisons T-6
- Grade Levels T-7

Other-School Comparison
- Description T-8
- Pros and Cons T-9

Same-School Comparison
- Description T-10
- Baseline Score T-11, T-12

Choosing an Approach
- Which Approach is Best? T-13
- When Neither Approach is Feasible T-14

Evaluation

Time

60-75 minutes

Materials needed

overhead projector and screen
blank transparency sheets
markers
evaluation forms
optional: flip chart, marker board, and/or blackboard

The advice given in this guide was clarified on February 17-19, 1992 during the National Schoolwide Projects Meeting.

NOTE TO PRESENTER

- In the comparison school model, the schoolwide project **must** compare the aggregate achievement scores of Chapter 1 students with the aggregated achievement scores of **all** Chapter 1 students served in non schoolwide project schools in the same grades in that district.

- For projects serving Pre-K, K and 1 norm referenced test scores may **not** be used for schoolwide project accountability purposes.

- At the end of the ***first 3-year cycle***, the determination for the accountability requirement may be made for the full three year period *or* for the third year only. However, at the end of the ***second 3-year cycle***, the determination for the accountability requirement must be made for the full three year period.

WORKSHOP GOAL

Introduce the workshop by stating the goal.

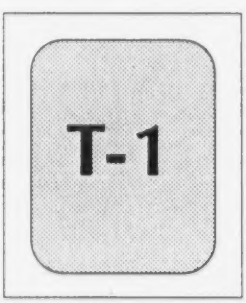
As a result of this workshop participants will learn about the special evaluation requirements for Chapter 1 schoolwide projects.

NOTE TO PRESENTER

This workshop may be presented as a stand-alone overview of the schoolwide project evaluation requirements for schools that have opted for this delivery model, OR it may be presented as a module in an all-day schoolwide project planning workshop.

Presenters are cautioned to stay close to the script. As this workshop presents current USED policy, deviation from the script may be viewed as interpretation, which we are strictly prevented from doing. It is also recommended that you review this workshop with a representative of the SEA to ensure that SEA policy coincides with USED policy.

OVERVIEW



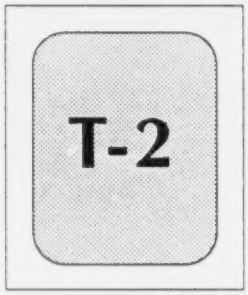
Display T-1:

"Schoolwide Projects Must Be More Effective." Explain that each schoolwide project is required to demonstrate that it is more effective than a traditional Chapter 1 service delivery model in improving the

achievement of its educationally disadvantaged students. (Otherwise, why have a schoolwide project?)

If greater effectiveness cannot be demonstrated at the end of the third year after the project was implemented, the school will lose its status as a schoolwide project and will revert to a traditional project.

OPTIONS FOR DEMONSTRATING EFFECTIVENESS



Display T-2:

"Options for Demonstrating Effectiveness." Demonstration of effectiveness may take one of two forms, to be specified in the schoolwide project application.

NOTE TO PRESENTER

Some SEAs do not require a choice to be made in advance; some do. Some allow schools to change the model to be used during the third year of implementation; some do not.

The two options are:

Same-School Comparison. Schoolwide project evaluation results in the aggregate must exceed aggregated results of Chapter 1 students from that school who were served prior to implementation of the schoolwide project.

Other-School Comparison. Schoolwide project evaluation results in the aggregate must exceed results of a comparable group of Chapter 1 students in the district as a whole.

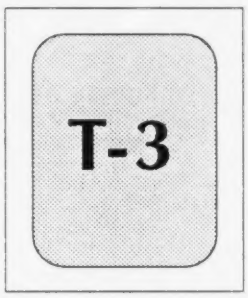
Either the mean or the median can be used as a measure of central tendency (average), but the metric to be used must be used consistently throughout. The median may be the best choice for a small program with a few outlier scores in either direction. For large programs, there is usually little or no difference between the mean and the median scores.

NOTE TO PRESENTER

If participants do not know the differences between mean and median, you may want to use an additional transparency illustrating the differences and cases in which they are meaningful.

Secondary schools may demonstrate accountability with lowered dropout rates, decreased retention rates, or increased graduation rates instead of greater achievement gains as long as aggregated achievement is not lower that it had been during the past three years.

MEETING ACCOUNTABILITY REQUIREMENTS

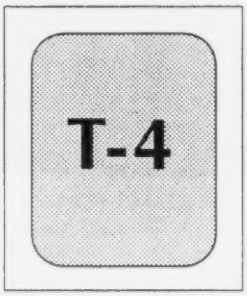


Display T-3:

"Key Questions." To initiate your schoolwide project evaluation study, consider the following questions:

1. Which subject areas must be included in an accountability study?
2. What measures of achievement may be used?
3. On which comparisons must the schoolwide project exceed the comparison groups?
4. What grade levels must be included?

Subject Areas



Display T-4:

"Subject Areas." To determine which subject areas should be included in the evaluation, ask the question, "What areas of instruction would be provided to these students through Chapter 1 in the absence of a schoolwide project?"

The targeting of subject areas must be supported through the needs assessment and reflected in the schoolwide project application.

Achievement Measures

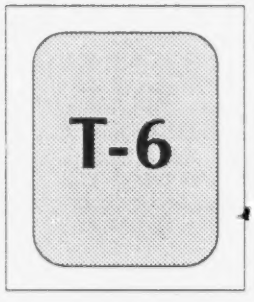
Display T-5:

"Achievement Measures." Measures of achievement must include, at a minimum, those tests used by the LEA to report aggregated performance for the Chapter 1 evaluation reporting system in the appropriate subject area(s). Both basic and more advanced skills scores must be compared. Comparisons must be based on annual gains on standardized achievement tests as measured by NCEs for grades two and above.

Additional measures may be used, subject to the approval of the SEA. They may include criterion-referenced tests, applied performance measures, and other objective measures of achievement. These measures are encouraged because they allow students to demonstrate their academic achievement in a variety of ways at different times. These measures may be reported as fall-spring gains.

Schoolwide project applications may be amended during the three years to specify additional measures to be used. These added measures must be approved by the SEA.

Comparisons



Display T-6:

"Comparisons." In projects that establish goals in more than one area of achievement or that use more than one achievement measure, the preponderance of evidence should favor the schoolwide project. That is, if several measures are used, the number of measures that reflect gains must exceed the number that do not reflect gains.

When comparisons involve the same metric such as NCE scores, the magnitude of the difference may be considered in determining whether the preponderance of evidence favors the schoolwide project.

For example, if the schoolwide project exceeds the comparison group by 4 NCEs in advanced skills in math but trails the comparison group by 1 NCE in basic skills in math, the positive comparison offsets the negative one. Comparisons may be made separately by grade level or aggregated across grade levels.

In many states, the LEA must identify the comparisons it will make in advance in its schoolwide project application. Comparisons must involve gains in basic and more advanced skills in at least one subject area (with the exception of the secondary school provision previously mentioned).

Grade Levels

Display T-7:

"Grade Levels." For accountability purposes, only grades served by Chapter 1 in the LEA as a whole or in the school prior to the initiation of the schoolwide project are to be compared. Prekindergarten, kindergarten, and first grade gains are to be computed separately from those in grades two and above.

OTHER-SCHOOL COMPARISON

Description

Display T-8:

"Other-School Comparison: Description." Using the other-school comparison model, gains on achievement tests of educationally disadvantaged children in the schoolwide project are compared with gains of educationally disadvantaged children in traditional projects in the same LEA. This comparison is valid only if similar groups of students are compared; that is, students in the same grades who are measured by the same tests and the same

testing cycles. (If the schools do not meet these criteria, they must use the same-school comparison option.)

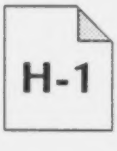
Pros and Cons

Display T-9:

"Other-School Comparison: Pros and Cons." An advantage of the other-school comparison model is the ability of a school to change tests and/or to add additional measures without much difficulty, as long as both the project school and the comparison schools make the same changes and the proper procedures are followed for score conversion.

One disadvantage of this model is the difficulty of finding schools that are truly comparable.

Another disadvantage is that the regression effect may work to the advantage of the comparison schools if those schools have higher-achieving populations than the project school. (The mean scores in such schools are higher, and regression to the mean, which occurs in every school, benefits the school with the higher mean.) There is a formula you may use to minimize this effect (see *H-1:* Davis and Quilling, pp. 8-9).



SAME-SCHOOL COMPARISON

Description

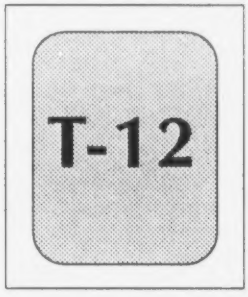
Display T-10:

"Same-School Comparison: Description." In same-school comparison, the gains of educationally disadvantaged students in the school must exceed the gains of educationally disadvantaged students in the same school during the three years prior to the start of the schoolwide project. The outcome measures, testing intervals, and demographic characteristics of the student population should remain constant for six years.

Baseline Score

Display T-11:

"Same-School Comparison: Compute Baseline Score/ Average Gains." The three-year baseline score must be computed by averaging annual gains for students in the three years prior to the initiation of the project in both basic and more advanced skills. You cannot compute a baseline score by averaging fall-spring gains.



Display T-12:

"Same-School Comparison: Compare Baseline Score/ Choices." Compare baseline gains only to the gains of the students who would be eligible for Chapter 1 in the absence of the schoolwide project. Comparisons can be made on the schoolwide project's three-year average OR on the schoolwide project's third-year gain only. (The rationale for using all three years is that you "wash out" the effects of any single year that may be problematic; i.e., where test scores are lower due to extraneous factors. The rationale for using the third year only is that it is the culmination of the schoolwide project and should represent full implementation.)

Change of tests:

Three-year average: You must equate the tests through the use of an equating table or by double-testing one year.

Third year only: Equating tests is not necessary (unless you change during the third year). The assumption is that NCE gains on nationally normed tests are comparable from one test to another.

CHOOSING AN APPROACH

Which Approach Is Best?

Display T-13:

"Which Approach Works Best?" Determining the best accountability approach depends on school circumstances:

Are comparable schools available?

Have there been changes in tests, test intervals, student population (from desegregation or mobility, for example), student retention rates, targeted grades, and so on?

The key issues are validity and effort.

The other-school comparison model is preferable when circumstances in the LEA have changed appreciably, when retrieving data is difficult, or when past measurements are considered invalid due to poor test-curriculum alignment or other problems.

The same-school comparison model is preferable when available comparison schools are unavailable or too different or when a school would like to motivate staff to raise achievement more than in the past. It also avoids possibly unpleasant comparisons to others.

When Neither Approach is Feasible

Display T-14:

"When Neither Approach is Feasible." When the SEA and LEA agree that neither approach is feasible--e.g., when there are no comparison schools or there is no way to retrieve "old" data, or when the school was formerly a schoolwide project--the accountability requirement is accomplished through the same evaluation procedure used to determine whether the school must develop a plan for program improvement. That is, the school must show substantial progress toward meeting desired outcomes and improvement in the aggregate performance of educationally disadvantaged children in the school over a three-year period. This determination is made at the end of the three-year period. Progress must be shown in both basic and more advanced skills. The schoolwide project is still responsible for demonstrating that the preponderance of evidence favors it. The project may use either the three-year average or third-year-only results as evidence.

In summary, the basic rules are:

1. You may use a same-school or other-school comparison model. Choose the model that is most valid and requires the least amount of effort.

2. The preponderance of evidence must favor the schoolwide project if the school is to continue as a schoolwide project.

Discuss:

"What makes the most sense for you?"

(optional)

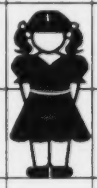
Present examples from Davis and Quilling.

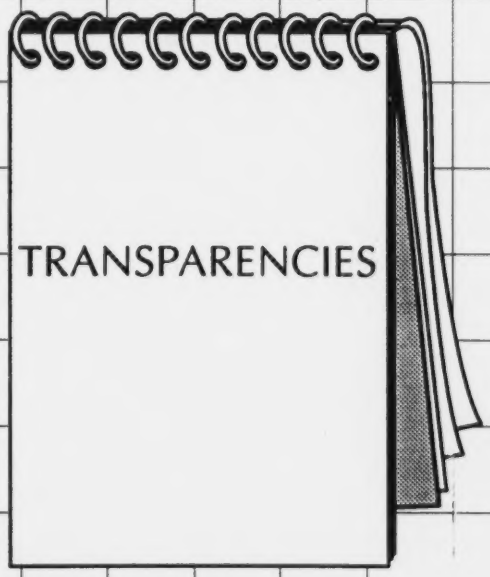
EVALUATION

Distribute the standard workshop evaluation forms.

WORKSHOP GUIDE

SCHOOLWIDE PROJECT EVALUATIONS





prepared by:

RMC RESEARCH CORPORATION
1512 Larimer Street, Suite 540
Denver, Colorado
80202

SCHOOLWIDE PROJECTS MUST BE MORE EFFECTIVE THAN TRADITIONAL CHAPTER 1 PROJECTS.

OPTIONS FOR DEMONSTRATING EFFECTIVENESS:

- same-school comparison
- other-school comparison.

KEY QUESTIONS:

- Which subject areas?
- What measures?
- Which comparisons?
- What grade level?

T-3

SUBJECT AREAS:

- those provided to students in the absence of a schoolwide project
- those supported by the needs assessment.

ACHIEVEMENT MEASURES:

- NRTs: basic/more advanced skills
- desired outcomes
- additional measures.

COMPARISONS:

Demonstrate the preponderance of evidence in favor of the schoolwide project.

- magnitude
- number

GRADE LEVELS:

- those served by the LEA as a whole

OR

- those served prior to the schoolwide project implementation

(Separate grades 2-12 from Pre-K, K and Grade 1)

T-7

OTHER-SCHOOL COMPARISON: DESCRIPTION

Schoolwide gains > other projects' gains

- similar students (identified using same criteria and same grade levels)
- same test, same testing cycle

OTHER-SCHOOL COMPARISON: PROS AND CONS

+ Can change tests/add measures
- Regression effect

SAME-SCHOOL COMPARISON: DESCRIPTION

Schoolwide gains > gains of Chapter 1 students served during previous three years

- same outcome measures
- same testing intervals (annual)
- same student demographics

SAME-SCHOOL COMPARISON:

- compute baseline score by averaging annual gains for three years in both basic and more advanced skills.

T-11

SAME-SCHOOL COMPARISON:

- compare baseline only to scores of Chapter 1 eligible children using either three-year average OR third year only.

WHICH APPROACH WORKS BEST FOR YOU?

- most valid
- least effort

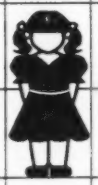
T-13

WHEN NEITHER APPROACH IS FEASIBLE:

- both SEA and LEA must agree
- use the same evaluation as for local annual review
- preponderance of evidence must favor the schoolwide project
- three-year average vs. third-year-only option applies.

WORKSHOP GUIDE

SCHOOLWIDE PROJECT EVALUATIONS



prepared by:

RMC RESEARCH CORPORATION
1512 Larimer Street, Suite 540
Denver, Colorado
80202

BEST COPY AVAILABLE

MEETING THE ACCOUNTABILITY REQUIREMENT

FOR SCHOOLWIDE PROJECTS

Alan Davis

RMC Research Corporation

and

Mary R. Quilling

PRC Inc.

February, 1992

Meeting the Accountability Requirement for Schoolwide Projects

Alan Davis and Mary R. Quilling

Introduction

Schoolwide projects may use Chapter 1 resources to improve the entire educational program of a school. In this respect they differ from other Chapter 1 projects in which resources are used to provide additional services only to students identified as educationally disadvantaged. It is possible that educationally disadvantaged students, who are the focus of P.L. 100-297, will receive less direct educational benefit in a schoolwide project if the project is not properly designed and implemented. To protect against this, schoolwide projects are subject to an additional accountability requirement: each schoolwide project is required to demonstrate that it is more effective than a traditional Chapter 1 program in improving the achievement of its educationally disadvantaged students. If greater effectiveness cannot be demonstrated after a three year period, a school will lose its status as a schoolwide project and revert to a traditional project.

This paper discusses the options available to schools to satisfy the accountability requirement for schoolwide projects. Its purpose is to provide general practical guidance in selecting and implementing an approach that will fit local circumstances. Local education agencies are encouraged to contact their Chapter 1 Technical Assistance Center and state education agency for more concrete assistance in planning and conducting their accountability studies.

Two Approaches to Demonstrating Accountability

The law (PL 100-297, Sec. 1015.e) and regulations (Federal Register, May 19, 1989, 21765-6) offer two approaches to meeting the accountability requirement for schoolwide projects. At the end of the third project year, the average [1] achievement gains of educationally disadvantaged students in the school must be compared to the gains of:

(a) a comparable group of Chapter 1 children in the district as a whole; or

(b) Chapter 1 children in the school during the three years prior to implementation of the schoolwide project.

---

1 Either the mean or the median can be used. The method must be specified in advance and used consistently throughout.

In this paper, we refer to the first approach as the Other Schools Comparison, and to the second as the Same School Comparison.

In either comparison, the gains of educationally disadvantaged students in the schoolwide project must exceed the gains of the comparable students. Secondary schoolwide projects may demonstrate accountability with lower dropout rates, decreased retention rates, or increased graduation rates instead of greater achievement gains as long as achievement levels during the three schoolwide project years are not lower than those of the three years preceding the schoolwide project.

We will discuss each approach in detail. Before we do, we will address some preliminary questions common to both approaches:

- What subject areas must be included in an accountability study?
- What measures of achievement may be used?
- On what comparisons must the schoolwide project exceed the comparison group?
- What grades must be included?

Each of these questions will be discussed in turn.

Subject Area Focus

In comparing the achievement gains of educationally disadvantaged students in schoolwide projects with gains of the comparison group, in what areas of achievement should the comparison be made?

Simply stated, the purpose of the accountability requirement is to determine whether students are learning more in a schoolwide project than they would in a Chapter 1 project targeted at individual students. To determine what areas of achievement should be included, one must ask: What areas of instruction would have been provided to these students through Chapter 1 in the absence of a schoolwide project? In keeping with Chapter 1 regulations, the targeting of subject areas must be supported by a needs assessment, and reflected in the application for the schoolwide project.

Example 1

The student needs assessment in Central School District has indicated a need for Chapter 1 instruction in both reading and mathematics. Chapter 1 services in non-schoolwide project schools are being provided in both areas. When Taft Elementary became a schoolwide project, the staff wanted to set a schoolwide goal to improve reading, but did not set a goal for math. Both math and reading services were provided by the district to other Chapter 1 schools during the first year of the schoolwide project. Is the schoolwide project accountable for improved gains in both reading and math?

Yes. In the absence of the schoolwide project, educationally disadvantaged students at Taft would have received Chapter 1 supported instruction in reading and in math, so the school is accountable for gains in both subjects.

Selection of Outcome Measures

What kinds of measures should be included?

The LEA must designate one or more valid and reliable measures of student achievement in basic and advanced skills in its application that it will use to determine whether the accountability requirement has been met. At a minimum, measures should include those tests used by the LEA to report aggregated performance for the Chapter 1 evaluation and reporting system in the appropriate subject(s). Additional measures may be used subject to the approval of the SEA. These may include criterion or domain referenced tests, applied performance measures, and other measures of achievement that are objective in the sense that they are administered and scored using procedures to minimize effects of rater expectation or bias.

Multiple choice measures are not developmentally appropriate for use below the second grade, and should not be used in those grades. At a minimum, measures should be selected for use below grade 2 that reflect the desired outcomes stated in the district Chapter 1 application at those levels.

Once an LEA has specified accountability measures in the plan for the schoolwide project, it must collect data and report the results. On the other hand, it may amend its application at any time to specify additional measures to be used in the future. Comparisons based on all measures are subject to the rule that a comparison favoring the comparison group on one measure must be more than offset by a comparison favoring the schoolwide project on another, as discussed later in this paper.

LEAs are encouraged to use multiple measures of achievement in meeting the accountability requirement. Relying on a single measure when important decisions rest on the outcome can encourage teaching narrowly to the format and content of the test. The use of several measures allows students to demonstrate their

achievement in a variety of ways and at different points in time. The use of applied performance measures, such as writing samples and math problem-solving demonstrations, is encouraged because such performances promote thoughtful instruction and require the integration and application of advanced skills.

Comparisons for grades 2 and above involving normed measures must be based on gains, and not on a single administration of the measure (Chapter 1 Regulations, May 1989, Section 200.36.f). Gains must be measured over a 12 month period. Comparisons involving grade-specific measures that are not vertically equated and normed (such as local writing samples or kindergarten performance scales) may be based on fall to spring gains or simple post-test comparisons.

Example 2

Urban School District provides Chapter 1 instruction in reading and language arts in grades 1-5. Students in grade 1 are assessed using an informal reading inventory and a performance checklist of language skills. The district administers a nationally normed achievement test battery to all students in grades 2 - 12 each spring. Gains on the Reading Comprehension and Total Language tests are aggregated and reported through the national Chapter 1 and evaluation reporting system. The district also administers a writing sample in the spring of grades 3 and 7 and a state minimum competency test in the spring of grade 6. In addition, students complete a self concept inventory in grade 5. The LEA would like to include all of these measures in meeting the accountability requirement for the schoolwide Chapter 1 project at Central Elementary.

Discussion: The nationally normed achievement test must be included, and expressed in gains. The self concept inventory cannot be included, because it is not a measure of achievement. The writing sample and the state minimum competency test are appropriate measures, since they correspond to the language arts goal of Chapter 1 in the district. The informal reading inventory and performance checklist used at grade 1 are appropriate if care is taken to ensure that they are reliably administered and scored. Since gains are not available for these measures, comparisons will be based on posttest scores. (Note that use of these measures requires their aggregation for Chapter 1 students in non-schoolwide projects as well.)

Criteria for Successful Comparisons

When goals include more than one area of achievement, or when more than one achievement measure is used, on what comparisons must the schoolwide project exceed the comparison group?

If a schoolwide project establishes goals in more than one area, or if several types of measures are employed, several comparisons may be involved. For example, a

project that sets out to improve achievement in reading and math in both basic and advanced skills could involve a minimum of four comparisons.

Since each comparison involves some amount of error, it is quite possible that a schoolwide project that is truly effective will not exceed the comparison group on every single comparison. It is sufficient that the "preponderance of the evidence" favor the schoolwide project. To demonstrate this, any comparison in which the comparison group exceeds the schoolwide project must be more than offset by a comparison favoring the schoolwide project.

When comparisons involve the same standard metric (such as NCE scores), the magnitude of the difference may be considered in determining whether a positive comparison offsets a negative comparison. For example, if the schoolwide project exceeds the comparison group by 4 NCEs in advanced skills in math but trails the comparison group by 1 NCE in basic skills in math, the positive comparison more than offsets the negative one.

When different metrics are used, the count of positive comparisons must exceed the count of negative comparisons. For example, if the schoolwide project exceeds the control group in average improvement on a local writing sample, but trails the comparison group in improvement on a state basic skills test and on a nationally normed achievement test, then the schoolwide project has failed to exceed the comparison group on the majority of comparisons.

The LEA must identify the comparisons it will make in advance in the application for the schoolwide project. With the exception of secondary schools, comparisons must involve gains in basic skills and more advanced skills in at least one subject area. Comparisons may be made separately by grade level or aggregated across grade level.

Example 3

Urban School District stated in its application that the accountability requirement for the schoolwide project at Central would be based upon annual gains on a nationally-normed test in Total Language Arts and Reading Comprehension, and on fall-spring gains on a writing sample given in grade 3 only. Because the writing sample was given at only one grade, comparisons were made for each grade separately. The results are shown below:

Average Gains
Of Educationally Disadvantaged Students

	Tot. Language (NCE Gain)		Reading Comp. (NCE Gain)		Writing (Av. Gain)	
Grade	Central	Comparison	Central	Comparison	Central	Comparison
2	2.8	3.1	3.6	3.5		
3	1.2	2.2	-0.8	2.1	0.9	-0.4
4	4.3	2.7	3.3	-1.0		
5	5.1	3.7	2.1	3.2		
6	2.7	1.4	2.9	1.9		

The table presents 11 comparisons between the schoolwide project and the comparison group: 10 involving NCE gains on nationally normed tests, and one involving average gains on a holistically scored writing sample. Of these 11 comparisons, 7 (underlined) favored the schoolwide project. Since the majority of comparisons favor the schoolwide project, the accountability requirement is satisfied, and the project is eligible for renewal.

Grades to Include

What grade levels must be included in the comparison?

A schoolwide project must be designed to upgrade the educational program in all the grades of the school (Regulatory Comment, Federal Register, May 19, 1989, p. 21792). However, the comparisons required by the accountability provision apply only to those grades served by Chapter 1 in the local education agency as a whole or in the school prior to the initiation of the schoolwide project. If students in a particular grade in the schoolwide project would be served in Chapter 1 in the absence of the schoolwide project, then that grade must be included in the accountability comparison. In grades below the second grade, comparisons should not be made on the basis of multiple-choice tests, but should reflect the desired outcomes stated in the district application for services at this level.

Example 4

Rural County School District provides Chapter 1 instruction to students in grades K-4. Harrison Elementary School, a K-5 school, has a schoolwide Chapter 1 project. To meet the accountability requirement, Rural County will conduct grade by grade comparisons of student achievement gains for educationally disadvantaged students in grades K-4. In grades K and 1, Rural County proposes to give individually administered tests of receptive and expressive language in the fall and the spring, and report gains in the total raw score. In grades 2-4, Rural County will use annual NCE gains on nationally normed tests. Grade 5 will not be included in the comparison, because fifth graders would not receive Chapter 1 services in the absence of the schoolwide project.

The Other Schools Comparison

The Other Schools Comparison involves comparison of the gains of educationally disadvantaged students in the schoolwide project with gains of a comparable group of Chapter 1 children served in traditional projects. Since the purpose of the comparison is to determine whether a schoolwide project is more effective than traditional projects, the comparison group must be made up only of students in traditional Chapter 1 projects, and should not include students in other schoolwide projects.

The comparison will be valid only to the extent that similar groups of students are compared. At a minimum, this means that the comparison of gains will be for students (a) identified by the same criteria and (b) in the same grades. The same tests and testing cycles must be used for both groups. If no such comparison group exists within the local educational agency, the Same School Comparison must be used instead.

A significant advantage of the Other Schools Comparison is that it allows the LEA to change tests or add additional measures without undue difficulty. LEAs can change tests or norms as often as they like, so long as the change affects the schoolwide project and comparison schools the same, and procedures are followed to allow the computation of gains during the year of the change. The Same School Comparison, on the other hand, presents a difficult problem when the LEA wants to change from whatever measures were used in the initial baseline period.

The primary disadvantage of the Other Schools Comparison is that other schools may present an unfair comparison because of differences in the students they serve. Because schools hosting schoolwide projects must be schools serving a high concentration of students from low-income homes, these schools are generally more likely to face environmental obstacles to learning, such as lower attendance, less parent involvement, and higher mobility. Schools serving low income areas may have a harder time attracting top faculty whose experience and prestige allow

them mobility within the system. These characteristics make it harder to raise achievement.

The effect of statistical regression on achievement scores may also work to the detriment of the schoolwide project. More affluent schools are generally higher achieving schools. If a district-wide cutoff on the pretest score is used to identify Chapter 1 students, identified students in a higher achieving school will deviate more from the mean of the group from which selection takes place than will identified students in a low-achieving school. Consequently, when a uniform pretest criterion is used to identify educationally disadvantaged students, regression to the mean is likely to contribute more to the gains of the comparison group than to those of the schoolwide project participants.

Selecting Students in the Schoolwide Project for Comparison

Differential regression, described above, can have a substantial effect on the outcome of the Other Schools Comparison. The problem is the same as that encountered by LEAs who used Model B of the older Title I Evaluation and Reporting System (TIERS), which involved comparing the gains of Title I students to the gains of comparable students in non-Title I schools. Although it would appear reasonable to identify comparable students by matching their scores or by applying the same cutoff score in both sets of schools, these approaches result in non-equivalent groups when the school populations themselves are not equivalent.

The regression problem can be minimized by (a) selecting students on the basis of measures other than the pretest, or (b) selecting groups so that they represent the same proportion of the larger pool of students from which they were selected. [1] To use the proportional approach, find the number of students served in Chapter 1 in a given subject area in the LEA by grade (not including schoolwide projects). Divide that by the number of students in each grade in Chapter 1 schools where services in that subject are provided (not including schoolwide projects). This procedure yields the proportion of students in Chapter 1 in each grade for a subject area.

Next, find the number of students in each grade of the schoolwide project and multiply it by the corresponding proportion found above. This number represents the number of students to be included in the accountability study for this grade. Rank students by the same types of scores used to select students in other Chapter 1 schools, omitting those who would not be eligible for Chapter 1 services on other grounds (e.g., those already receiving comparable services), up to the number required for that grade. Repeat this procedure for each grade and subject area. Note that this procedure will not yield the same students identified as educationally disadvantaged for purposes of funding or annual evaluation. It will,

1 This was precisely the approach used in TIERS Model B. See Talmadge, K., Wood, C., and Gamel N. (1981). User's Guide: ESEA Title I Program Evaluation, p. 47. Washington, DC: US Department of Education.

however, yield a sample for the accountability comparison that is less biased against the schoolwide project.

Example 5

Adams Elementary School, a K-5 school, has a schoolwide project. There are 10 other Chapter 1 schools in the LEA. Chapter 1 reading instruction is provided in each of these 10 schools in grades K-3. In addition, four of the schools offer math instruction through Chapter 1 in grades 2-5.

The LEA uses different criteria to select students for Chapter 1 in reading and math. In reading, students must score below the 35th percentile on a nationally normed test (used as the pretest), and be rated "significantly low" in classroom performance by the classroom teacher. In math, students must score below the level of minimal acceptable performance on a math test developed by the LEA (not used as a pretest).

To use the Other Schools Comparison for reading, the LEA will first identify the proportion of students served in Chapter 1 reading in the other schools, and then identify a corresponding proportion of students within Adams Elementary who would be served in Chapter 1 reading and math in the absence of the schoolwide project. The proportions for the LEA in reading are shown below:

Chapter 1 Reading Enrollment as a Proportion of Total Enrollment of Chapter 1 Schools

Grade	Total N	Chap 1 Reading	Percent Reading
K	621	128	20.6%
1	638	142	22.2%
2	617	186	30.2%
3	598	144	24.0%

Now, the students in Adams must be ranked in order of need using the same selection criteria used in other schools. In reading, these include teacher judgments of student performance and scores on nationally normed tests. Students who would not be served in Chapter 1 for other reasons will be omitted. The first 20.6% of Kindergarten students listed, the first 22.2% of first graders, and so on, will constitute the comparison group for reading. In math, it is not necessary to use this proportional procedure because students are not selected on the pretest and regression to the mean is not a serious problem.

Third Year vs. All Three Years

The comparison is to be made at the end of the three year period. If a schoolwide project is new, in its first three-year cycle, the comparison at the end of the third year may be for all gains during the three year period or between gains during the third year only. Taking into account the imperfect reliability of gains in any given year (especially in small projects), including gains for all three years generally provides a more stable indicator of impact. On the other hand, if the project improves during the three year period, it should be able to rely upon the third year results. If the project has already completed its first three-year cycle, the comparison must include all gains during subsequent three year periods.

The Same School Comparison

A schoolwide project may satisfy the accountability requirement by demonstrating after three years that the achievement gains of educationally disadvantaged students in the school exceed the average achievement gains of comparable educationally disadvantaged students in the same school in the three years prior to the start of the schoolwide project.

This option is appealing when there are problems finding equivalent comparison groups within a district. However, constructing equivalent historical comparison groups often poses significant difficulties that must be considered carefully. Ideally, to conduct the Same School Comparison, the school should remain relatively constant for six years in respect to:

- the outcome measures administered,
- the testing interval used, and
- the demographic characteristics of the student population.

Establishing the 3-Year Baseline

Before the passage of P.L. 100-297, most Chapter 1 projects measured achievement gains from fall to spring, and did not report gains for basic and advanced skills separately. Schools that established schoolwide projects soon after the passage of the new legislation may find it difficult or impossible to establish a baseline of 12-month gains in basic and advanced skills against which to compare subsequent gains. To do so in most cases would require re-examining the scores of each individual student, determining gains in basic and advanced skills on a 12-month testing cycle, and then computing new average gains for each grade for each of three years before the initiation of the schoolwide project. Note that simply using the average scores of two spring testing points will not accomplish the same result, because the averages will not be restricted to those students with both pre and posttest scores. Note also that should the LEA compare annual gains of the

schoolwide project to fall-spring gains in previous years, the schoolwide project is very likely to fail in the comparison because gains measured fall-spring tend to be higher.

Example 6

Adams Elementary School became a schoolwide project at the same time that the LEA moved from fall-spring to annual testing in Chapter 1. Adams is very different from other Chapter 1 schools in the LEA, and elects to be compared to its previous performance rather than to other schools. But since future testing will be done only once a year, previous fall-spring gains are not a fair comparison.

To re-calculate annual gains, the LEA Chapter 1 coordinator worked with a clerk to locate the scores of Chapter 1 students for the three previous years. Calling the first of the three years "Year 1," they found the spring scores of Chapter 1 students in both basic and advanced skills for that year. They then searched testing records for the spring of the previous year to find matching pretest scores for these students. Matching scores were located for about 70% of the students. Averaging the two sets of spring scores, they calculated the average spring-spring gain for Year 1, and repeated the process for Year 2 and Year 3. Then, for each grade level they averaged together Year 1, Year 2, and Year 3 to find the average for the baseline period. Those averages became the standard for their Same School Comparison.

Identifying Students for Comparison

A same-school comparison will be valid only to the extent that equivalent groups are compared across time. To accomplish this, the schoolwide project must identify students who would have been served in Chapter 1 in that school in the absence of a schoolwide project. It must employ the same procedures for selection used during the previous three years, selecting students in the same grades as in the past, using the same selection criteria.

If changes occurred during the baseline period in the grades included in the project or the selection criteria employed, then both comparison sets should be re-constructed employing the most restrictive service parameters.

Example 7

Brook School, a K-6 school, served students scoring below the 30th percentile in grades 1-5 in a Chapter 1 reading program for two of the three years before becoming a schoolwide project. The year before becoming a schoolwide project, it dropped grade 5, and continued serving 1-4. For the past three years, the school has operated as a schoolwide project with an objective to improve reading in all grades, and has considered all students below the 50th percentile as educationally disadvantaged. Using the most restrictive parameters, the comparison for School A will include only students identified as below the 30th percentile in reading in grades 1-4 both before and after becoming a schoolwide project.

Third Year vs. Three-Year Average

Gains for the three years of the baseline period are averaged together to establish the criterion for the Same School Comparison. After three years, the gains of educationally disadvantaged students in the schoolwide project will be compared to this criterion. If the project is new, in its first three-year cycle, the project may choose to compare the baseline average to the gains of its students in the third year only, or to the average of gains across the three years of the project. In subsequent cycles, the comparison must include all gains during the three year period.

Example 8

Taft School provided Chapter 1 services in reading in grades K-3 before becoming a schoolwide project. Taft chose to use the Same School Comparison to satisfy the accountability requirement. During the three years preceding the schoolwide project, gains in grades K and 1 were reported in increases in the number of skills mastered. In grades 2 and 3, annual NCE gains were reported in both basic and advanced skills. Results for the baseline period and for the first year of the project are shown below. To simplify the table, only gains in basic skills are displayed:

Average Gains for Educationally Disadvantaged Students

	Baseline Year				Project Year
Grade	1	2	3	Average	1
K	18.1	17.3	19.2	18.2	18.6
1	12.1	15.2	13.7	13.7	13.5
2	1.6	-1.2	2.2	0.9	1.4
3	3.3	4.1	1.7	3.0	2.2

Since the gains in different grades involve different types of scores, gains for all four grades cannot be averaged together. Instead, comparisons are made grade by grade. From the table, it can be seen that gains in the first year of the schoolwide project exceed the baseline average in grades K and 2, but not in grades 1 and 3. Since the first year involved major changes in grouping and instruction, gains in subsequent years are expected to be higher. The school may base its comparison entirely on the third year of the project.

Changes in Tests

The Same School Comparison is most valid when gains are based upon the same tests (or equivalent forms of the same tests) throughout the six year comparison period. On the other hand, LEAs will continue to select new measures, and publishers of norm-referenced tests will provide new norms. Use of the Same School Comparison cannot lock in testing practices indefinitely.

The year a new test is adopted poses a special problem. The LEA may use equating tables from a study satisfying Chapter 1 standards for test equating to adjust pre-test scores for the year of adoption of the new test. If no such study has been conducted, then students in the schoolwide project must be tested with both the old and new instrument during the adoption year for the new test if the project wishes to retain the option of averaging gains over the three project years. If only the gains from the third year are to be used for the accountability comparison, then a change in tests is simple, so long as it does not occur during the third project year.

If the new test and the previous test are nationally normed, then gains for subsequent years can be used without adjustment, based on the assumption that NCE gains are comparable from one test to another.

Example 9

After administering the Old Standard Achievement Test for 5 years, Central School District changed to the New Improved Test from a different publisher. The change occurred during the second year of a new schoolwide project. No equating study was available from either publisher. The year of the change, the LEA administered the new test to all students in the spring. It also administered the old test to those students in the schoolwide project who were identified as educationally disadvantaged for purposes of the Same School Comparison study and who had a pretest from the previous spring. Gains for Year 2 were based on the old test; the new test given that spring became pretest scores for Year 3. By following this procedure, the LEA left itself the option of using the average of three years of gains for the schoolwide project, rather than relying exclusively on the gains for Year 3 using the new test.

Selecting the Accountability Approach

For many schoolwide projects, circumstances will dictate the choice of accountability approach. Schoolwide projects in sparsely populated areas or in educational agencies in which all schools are schoolwide projects may have to use the Same School approach because no comparable traditional Chapter 1 projects are available as a comparison group. Other schoolwide projects without a three-year history of Chapter 1 must compare themselves to other schools. For most, the choice involves weighing considerations of validity and effort.

Because it uses current data, the Other Schools Comparison is preferable when circumstances in the LEA have changed appreciably over the six year period. This approach is unaffected by changes in tests, testing intervals, and data editing procedures when gains for the third year are compared. The Same School Comparison, on the other hand, is compromised by changes in the population served by the school due to desegregation policies or shifts in boundaries of school attendance areas. Differences in student selection procedures, such as inclusion of LEP students or service to retained students may also weaken the Same School Comparison.

Both approaches may involve time-consuming re-calculation of gains for the school-wide project to insure equivalence of comparison groups. However, these efforts are generally greater for the Same School Comparison, particularly when the school has changed its testing cycle, tests, selection criteria, or targeted grades during the six-year period. Retrieval of data 3 to 6 years old is likely to be a frustrating and time consuming task, particularly if the data are not already computerized. The cost of searching records by hand and the likely compromises in data quality are further cautions against a quick choice of the Same School approach.

The Same School Comparison has the advantage of helping to motivate staff by calling on everyone to raise achievement more than it has been raised in the past, and it avoids invidious comparisons with other current projects. The Same School Comparison may yield a more valid result when traditional schools in the district are too different from the school hosting the schoolwide project to provide an equivalent comparison group. Sooner or later, however, schools using this approach will face changes in tests or norms that may affect gains in ways that cannot be predicted. The approach also poses a constraint on those LEAs who would prefer to introduce very different types of assessments than those used in the past.

When Neither Approach is Feasible

There are local situations in which neither the Same School Comparison nor the Other Schools Comparison can be used. For example, consider a small rural education agency in which the only Chapter 1 school at each level has already operated as a schoolwide project under previous legislative provisions. It may well be that there is no baseline of data previous to the establishment of the schoolwide project for a Same School Comparison, yet there are no other Chapter 1 schools in the agency for the Other Schools Comparison. In this situation (which must be confirmed by the SEA), the accountability requirement is accomplished through the same evaluation procedure used to determine whether the school must develop a plan for program improvement.

The local education agency must demonstrate that the schoolwide project has shown substantial progress toward meeting the desired outcomes described in its application, and has shown improvement in the aggregate performance of educationally disadvantaged children in the school over a three year period. This determination is made yearly as part of the annual review, but the determination for the accountability requirement is made at the end of the three year period. At that time, the determination may be based on gains for the full three year period, or on gains during the third year only, using the "preponderance of evidence" approach described previously. The comparison must include both basic and advanced skills. If the evidence does not demonstrate both an improvement in aggregate performance and substantial progress toward meeting the desired outcomes described in the application, then the project will not be authorized to continue schoolwide.

WORKSHOP GUIDE

SCHOOLWIDE PROJECT EVALUATIONS

RESOURCE MATERIALS

Not included in the copy received by ERIC.

prepared by:

RMC RESEARCH CORPORATION
1512 Larimer Street, Suite 540
Denver, Colorado
80202